

基于 YOLOv5s 的轻量级海洋塑料垃圾检测

刘炳如¹ 万丽莉^{1*}

LIU Bingru WAN Lili

摘要

针对海洋塑料垃圾检测模型复杂度高且实时性差的问题,文章基于 YOLOv5s 提出了一种改进的轻量化目标检测模型 YOLOv5s-MSD。首先,采用 MobileNetV4 替换原 CSPDarknet53 主干网络降低模型复杂度;然后,采用 SDI 多层次特征融合模块替换 Concat 模块,通过整合编码器生成的层级特征图来增强图像中的语义信息和细节信息,从而增强模型的精确度;最后,结合轻量化模块 CSPPC 替换颈部 C3 模块,实现模型在轻量化和检测准确性上的双重提升。仿真实验表明,在海洋塑料检测数据集上, YOLOv5s-MSD 模型的准确率达到 96.8%。模型大小仅为 5.0 MB,参数数量和运算浮点数相较于原 YOLOv5s 分别减少 56.55% 和 74.68%。

关键词

海洋塑料垃圾;轻量化;目标检测; CSPPC

doi: 10.3969/j.issn.1672-9528.2026.03.014

0 引言

海洋污染,尤其是海底垃圾问题,已成为全球性环境挑战。根据联合国环境规划署的数据,全球每年有大约 800 万吨塑料进入海洋^[1]。塑料垃圾在 5 年内就能到达海洋最深处。海洋环境复杂多变导致海洋塑料垃圾人工清理困难,有效监测和管理海洋塑料垃圾变得尤为重要^[2-3]。周冬梅等^[4]提出了 WG-YOLO 算法,其 mAP50 达到了 68.1%, mAP50:95 达到 48.7%。尽管该算法在水下目标检测领域展现出一定的有

效性,但仍存在检测精度较低且模型不够轻量化的情况。针对上述问题,本文提出了一种基于 YOLOv5s 的轻量级海洋垃圾检测模型 YOLOv5s-MSL,该模型以复杂海洋环境中的塑料垃圾数据集训练而成,在保持高精度的同时实现轻量化,为海洋垃圾实时监测提供了可行方案。

1 YOLOv5s-MSD 模型

1.1 YOLOv5s-MSD 网络结构

YOLOv5s-MSD 网络结构如图 1 所示,主干网络为 MobileNetV4,其采用多尺度特征金字塔结构,包含五个层次的特征图(P1~P5),分别对应不同下采样率

1. 武汉轻工大学电气与电子工程学院 湖北武汉 430048

[基金项目] 武汉轻工大学科研项目资助(2025Y46)

参考文献:

- [1] 杜雨亭,何嘉.基于 YOLOv8 的雾天环境下车辆行人检测算法研究[J].软件导刊,2025,24(12):18-24.
- [2] 伍锡如,郝家琦,赵一波,等.复杂天气条件下基于 YOLO-CGT 的自动驾驶车辆检测[J].光学精密工程,2025,33(19):3135-3149.
- [3] 崔明义,冯治国,代建琴,等.基于深度可分离卷积和残差注意力模块的车道线检测方法[J].激光杂志,2024,45(4):81-87.
- [4] 马鸽,李洪伟,严梓维,等.基于多注意力的改进 YOLOv5s 小目标检测算法[J].工程科学学报,2024,46(9):1647-1658.
- [5] 牟毅,黄海松,李宜汀,等.改进 YOLOv11s 的无人机图

像小目标检测模型[J].电光与控制,2026,33(1):51-57.

- [6] 许牛琦,王金,付迪,等.基于 YOLOv11 改进的驾驶员视角下前方车头与车尾检测[J].信息技术与信息化,2025(11):106-109.

【作者简介】

宋梓棋(2006—),女,四川德阳人,本科在读,研究方向:深度学习、目标检测。

杨丹(1995—),通信作者(email:1253026013@qq.com),女,四川成都人,硕士研究生,教师、工程师,研究方向:深度学习、计算机视觉。

(收稿日期:2026-01-05 修回日期:2026-03-11)

P1 为浅层高分辨率特征图，保留了丰富的空间细节信息；P5 为深层低分辨率特征图，具有最大的感受野和高级语义信息。颈部网络中采用 SDI 模块替换 Concat 模块，采用 CSPPC 模块替换 C3 模块。

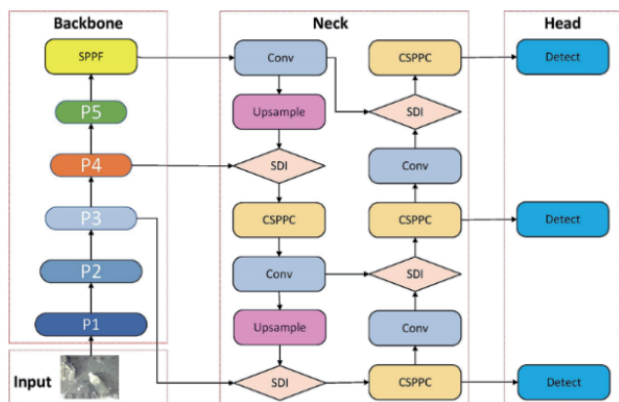


图1 YOLOv5s-MSc 网络结构图

1.2 MobileNetV4 主干网络

主干网络提取的多尺度特征为目标检测任务提供了丰富的图像信息，这种多尺度特征的生成和利用是 YOLOv5 实现高效且精确目标检测的关键。MobileNetV4 引入了多种新颖且高效的架构组件，其中最关键的是通用倒置瓶颈（UIB）。UIB 是对传统倒置瓶颈的进一步优化。

传统倒置瓶颈的数学表达式为：

$$y = \text{Proj}(\text{DWConv}(\text{Exp}(\mathbf{x}))) \quad (1)$$

式中：Proj 为投影层；DWConv 为深度可分离卷积；Exp 为扩展层； $\mathbf{x}$ 为输入特征图。

扩展层的作用是通过 1×1 卷积将输入的低维特征空间显式扩展至高维空间，深度可分离卷积是对扩展后的高维特征进行空间滤波，投影层的作用是通过 1×1 卷积将高维特征压缩回低维。

UIB 的数学表达式为：

$$y = T_p(\text{DWConv}(T_e(\mathbf{x}))) \quad (2)$$

式中： T_p 和 T_e 是动态线性变换层； $\mathbf{x}$ 是输入特征图。

UIB 通过引入可学习的动态线性变换层 T_p 和 T_e 替代传统固定倍数的通道扩展和压缩，使网络能够依据输入特征自动优化计算路径，提高了模型的信息利用率。

1.3 SDI 多层次特征融合模块

SDI 模块是 U-Netv2 的关键组件（如图 2 所示），通过整合编码器的多层次特征（ $L_1 \sim L_4$ ）增强语义与细节信息^[5]。该模块结合空间和通道注意力机制：空间注意力聚焦海洋塑料垃圾的潜在区域，抑制复杂水下环境（如光照不均、悬浮物干扰）导致的背景噪声；通道注意力优化特征通道的判别性，适应水下光照变化引起的颜色纹理差异。二者协同强化

了局部空间与全局通道特征的融合，显著提升了模型对模糊边界目标的表征能力。

每个层级通过空间和通道注意力机制后，不同层级的特征图通过不同的操作进行尺寸调整。较高级别的特征图（ $L=4$ ）经过上采样，使其尺寸与当前层级匹配，这样可以带入更多的细节信息；中间层级（ $L=3$ ）的特征图通过身份映射保持原尺寸传递；较低层级的特征图（ $L=2$ 和 $L=1$ ）通过下采样缩小尺寸，这样可以使其语义信息更好地融合到当前层级。这种多层次处理策略能够在保持特征图尺寸匹配的同时，增强不同层级之间的信息融合效果。为消除多层次特征融合时的空间分辨率与语义不一致性问题，SDI 模块采用平滑卷积对特征图进行优化处理。该操作通过卷积核的平滑作用有效减少层级间的不连续性，在保持细节完整性的同时抑制噪声干扰，从而生成连贯一致的特征表示。这种处理显著提升了模型对复杂水下环境中塑料垃圾的多尺度特征融合能力，最终输出优化后的统一特征用于下游任务。

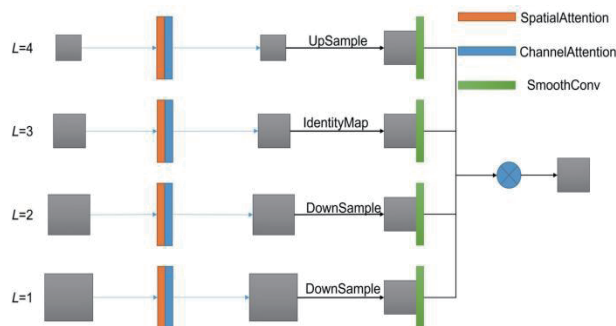


图2 SDI 网络结构图

1.4 CSPPC 轻量化模块

CSPPC 模块采用多分支结构，融合了双重卷积（double convolution）的多路特征提取能力，在保证信息丰富性的同时避免冗余。其轻量化设计核心为部分卷积（PConv），该结构仅在部分输入通道执行空间卷积，大幅降低计算量和内存访问开销。例如，1/4 部分率可减少计算至常规卷积的 1/16。PConv 通过保留关键通道并结合后续逐点卷积，维持了跨通道的特征流通性，在提升推理速度的同时保持特征提取精度，实现了高效轻量化的网络设计。

CSPPC 模块的核心思想是分支结构结合 PConv，来优化特征提取的效率和速度。CSPPC 的网络结构图如图 3 所示。CSPPC 模块从一个卷积模块开始进行特征图的处理。卷积核大小为 1，步长和填充分别设定为 1 和 0，输入特征图的通道数调整为 C_{out} 。卷积模块的作用是初步进行通道压缩和信息提取，为后续分支处理提供输入。接下来，特征图被分支结构分为两个分支，每个分支的通道数均为 $H \times W \times C_{out} \times 0.5$ ，其通道分离不仅减轻了计算负担，还为并行特征处理提供了可能性。其中一个分支通过多个 PConv 进行处理，PConv 模

块的重复次数可以根据任务需求灵活设定为 N 次。每次重复时特征图的通道数和尺寸保持不变，能够在不增加计算复杂度的前提下，层层递进地捕捉丰富的特征表示。两条分支通过 Concat 模块在通道维度上进行拼接，拼接后特征图的通道数为 $H \times W \times C_{out} \times 0.5 \times (N+2)$ ，这种特征融合方式保证了不同分支特征的有效整合。这样既保留了经过 PConv 处理的高效特征，又保留了未经处理分支中的细节信息。最后，拼接后的特征通过一个标准卷积层，输出特征图的通道数为 C_{out} ，既恢复通道数又输出标准特征图。

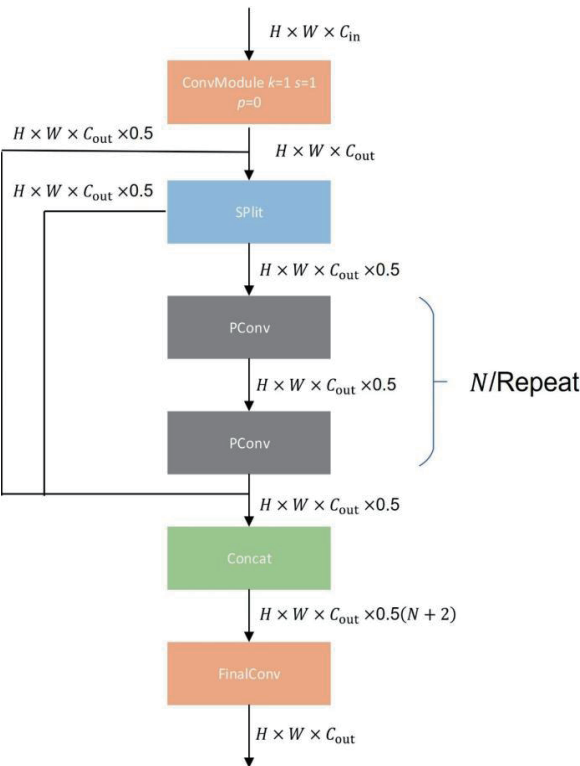


图 3 网络结构图

2 实验结果与分析

2.1 数据集

本研究采用 Kaggle 平台的 Underwater Trash Detection 数据集进行模型训练与测试，该数据集源自日本 JAMSTEC 开发的 J-EDI 深海影像档案，经处理后包含日本周边海域的水下图像。数据集标注了植物、动物及塑料垃圾（如塑料袋、饮料瓶等）三类标签，共计 7 667 张图像，其中训练集 6 133 张，测试集 1 534 张。

2.2 实验环境、参数及评价指标

所有实验平台在 Windows10 操作系统下进行，处理器型号为 AMD EPYC 7642 CPU，显卡为 NVIDIA RTX 3090。深度学习框架采用 PyTorch 1.12.1，编程平台为 PyCharm，编程语言为 Python 3.8。此外，还使用 Ubuntu 18.04 和 CUDA 11.3，实验环境如表 1 所示。

表 1 实验环境

配置	版本或型号
操作系统	Windows10
CPU	AMD EPYC 7642
GPU	NVIDIA RTX 3090
深度学习框架	PyTorch1.12.1
编程平台	PyCharm
编程语言	Python 3.8
Ubuntu	18.04
CUDA	11.3

实验采用随机梯度下降法 (SGD) 优化器训练模型，批次大小为 32，训练轮数为 100，图像输入大小为 640 px×640 px，初始学习率设置为 0.01，权重衰减系数设置为 0.000 5，并将动量设置为 0.937。未提及的其他超参数与 YOLOv5s 中的超参数保持一致。

实验中选取 mAP50、参数量、模型大小和浮点运算次数作为模型的评价指标。mAP50 表示 IoU 阈值为 0.5 时的平均精度的均值，参数量、模型大小和浮点运算次数都是衡量模型轻量化的评价指标。

2.3 消融实验

为验证三项改进对性能的提升效果，在相同的实验环境下进行消融实验，消融实验结果如表 2 所示。

表 2 消融实验结果

MobileNetV4	SDI	CSPPC	mAP50 /%	参数量 /10 ⁶	模型大小 Size /MB	浮点运算次数 /10 ⁹
			94.8	7.018	13.7	15.8
✓			95.5	3.282	6.7	4.4
✓	✓		96.3	3.275	6.7	4.5
✓	✓	✓	96.8	3.049	5.0	4.0

YOLOv5s 模型的主干网络替换为 MobileNetV4 后，mAP50 增加了 0.7%，参数量、模型大小和浮点运算次数分别为 3.282×10^6 、6.7 MB 和 4.4×10^9 ，与 YOLOv5s 相比分别下降了 52.23%、51.10% 和 72.15%。这表明，MobileNetV4 轻量化网络不仅更好地提取了特征信息，提高了准确度，还大幅度降低了模型参数量。基于以上改进策略，将颈部的 Concat 模块替换为 SDI，模型的 mAP50、参数量、模型大小和浮点运算次数分别为 96.3%、 3.275×10^6 、6.7 MB 和 4.5×10^9 。虽然模型大小没有变化并且浮点运算次数增加了 0.1，但是参数量有所下降。值得一提的是，mAP50 相比 95.5% 增加了 0.8%。最后，将轻量化模块 CSPPC 替换颈部网络中的 C3 模块后，mAP50 为 96.8%，相比于 96.3% 增加了 0.5%；参数量为 3.049×10^6 ，与 3.275×10^6 相比下降了 6.90%；模型大小

为 5.0 MB, 与 6.7 MB 相比下降了 25.37%; 浮点运算次数为 4.0×10^9 , 与 4.5×10^9 相比下降了 11.11%。这表明, 轻量化模块在面對质量较低图像时仍能做出准确又快速的反应, 在减少模型参数的同时, 模型准确度也有所提升。

2.4 对比实验

为进一步验证改进 YOLOv5s 模型的性能优势, 将该模型与 4 种主流的一阶段目标检测模型进行对比, 其中包括 YOLOX-tiny、YOLOv6n、YOLOv9n 和 YOLOv10n, 对比结果如表 3 所示。YOLOv5s-MS-C 的平均精度均值最高, 相较于 YOLOX-tiny、YOLOv6n、YOLOv10n 和 YOLOv11n 提升了 11.2%、2.1%、2.2% 和 3.5%, 浮点运算次数和模型大小均为最低。YOLOv10 整体稍逊色, 虽有较低的参数量, 但是浮点运算次数较高, 导致其在运行过程中计算效率降低、推理速度较慢、能耗与功耗的需求较大, 同时过多的卷积层可能会引发过拟合。改进后的 YOLOv5s-MS-C 模型在保持算法精度的基础上, 对模型的大小和浮点运算次数进行了较为明显的降低, 有助于在长续航的海洋监测装置上进行部署。

表 3 YOLOv5s-MS-C 与其他算法的对比实验

	mAP50 /%	参数量 / 10^6	模型大小 Size /MB	浮点运算 次数 / 10^9
YOLOX-tiny	85.6	5.033	10.1	7.6
YOLOv6n	94.7	4.427	5.4	11.5
YOLOv9n	94.6	2.699	5.2	7.2
YOLOv10n	94.7	2.427	5.4	8.5
YOLOv5s-MS-C	96.8	3.049	5.0	4.0

2.5 实验验证

为直观地显示改进模型的优势, 用 1 534 张测试集图像进行验证, 选取部分样本图片进行对比。图 4 (a) (c) (e) 为原模型检测结果, 图 4 (b) (d) (f) 为 YOLOv5s-MS-C 模型检测结果。

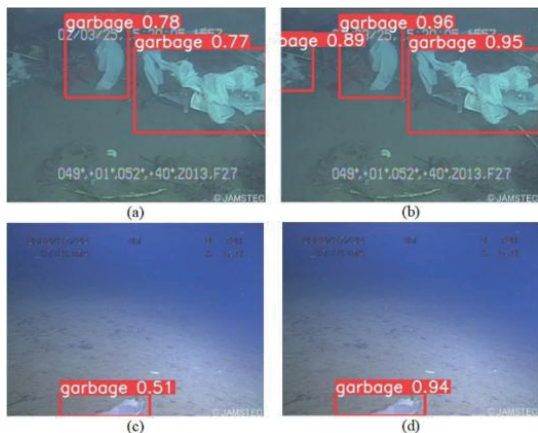


图 4 YOLOv5s 和 YOLOv5s-MS-C 的检测效果

通过图 4 (a) (b) 的对比, 在图片质量不佳的情况下, 能够将漏检的部分成功识别, 有效地提高了模型检测的性能。图 4 (c) (d) 的对比结果展示了改进模型在图片拍摄不全或有遮挡的情况下对目标能够更为精准地检测, 同时有效地避免出现误检和漏检的情况。图 4 (e) (f) 对比结果可知改进模型在检测小目标的任务上更为准确, 提升了检测准确率。

3 总结

本文对海洋塑料垃圾检测方法进行了研究, 以 YOLOv5s 模型为基准模型, 提出了一种精度高且轻量化的深度学习模型 YOLOv5s-MS-C。首先, 将主干网络更换为 MobileNetV4, 有效地降低了模型复杂度, 提高了模型的推理速度; 其次, 将颈部 Concat 模块替换为 SDI 模块, 这增加了模型检测海洋垃圾的准确性。最后, 融合 CSPPC 模块替换颈部 C3 模块, CSPPC 模块不仅减少了模型的计算量, 还增强了模型的特征提取能力。这表明本文所开发的轻量化算法可为海洋塑料垃圾检测领域提供有价值的参考。

参考文献:

- [1] 张雅清, 熊勇, 朱清. 基于改进 YOLOv5s 的水下自主机器人垃圾检测技术研究 [J]. 计算机测量与控制, 2025, 33(6): 76-85.
- [2] 张岩松, 高茂庭. 基于改进 YOLOv5s 的漂浮垃圾目标检测算法 [J]. 计算机工程与设计, 2025, 46(3): 910-917.
- [3] 庞梅, 汪珏, 詹泳, 等. 基于 YOLOv5 改进算法的海洋水下垃圾检测方法 [J]. 计算机与现代化, 2024(7):120-126.
- [4] 周冬梅, 王飞成, 吴小所. 改进 YOLOv8n 的海洋垃圾目标检测算法:WG-YOLO[J/OL]. 计算机工程与科学, 1-11[2026-02-24].<https://link.cnki.net/urlid/43.1258.TP.20250428.1343.004>.
- [5] 袁红春, 臧天祺. 基于注意力机制及 Ghost-YOLOv5 的水下垃圾目标检测 [J]. 环境工程, 2023,41(7):214-221.

【作者简介】

刘炳如 (2000—), 女, 山东德州人, 硕士研究生, 研究方向: 计算机视觉、深度学习。

万丽莉 (1983—), 通信作者 (email: wanlili@whpu.edu.cn), 女, 河南信阳人, 博士, 副教授、硕士生导师, 研究方向: 无人系统、智能控制。

(收稿日期: 2025-07-31 修回日期: 2026-03-06)